

Wright State University

CORE Scholar

---

International Symposium on Aviation  
Psychology - 2021

International Symposium on Aviation  
Psychology

---

5-1-2021

## Exploring Methods to Collect and Analyze Data on Human Contributions to Aviation Safety

Jon Holbrook

Follow this and additional works at: [https://corescholar.libraries.wright.edu/isap\\_2021](https://corescholar.libraries.wright.edu/isap_2021)



Part of the [Other Psychiatry and Psychology Commons](#)

---

### Repository Citation

Holbrook, J. (2021). Exploring Methods to Collect and Analyze Data on Human Contributions to Aviation Safety. *49th International Symposium on Aviation Psychology*, 110-115.  
[https://corescholar.libraries.wright.edu/isap\\_2021/19](https://corescholar.libraries.wright.edu/isap_2021/19)

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2021 by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

## EXPLORING METHODS TO COLLECT AND ANALYZE DATA ON HUMAN CONTRIBUTIONS TO AVIATION SAFETY

Jon Holbrook  
National Aeronautics and Space Administration  
Hampton, VA

Focusing on undesired operator behaviors is pervasive in system design and safety management cultures in aviation. This focus limits the data that are collected, the questions that are asked during data analysis, and therefore our understanding of what operators do in everyday work. Human performance represents a significant source of aviation safety data that includes both desired and undesired actions. When safety is characterized only in terms of errors and failures, the vast majority of human impacts on system safety and performance are ignored. The outcomes of safety data analyses dictate what is learned from those data, which in turn informs safety policies and safety-related decision making. When learning opportunities are systematically restricted by focusing only on rare failure events, not only do we learn less (and less often), but we can draw misleading conclusions by relying on a non-representative sample of human performance data. Changes in how we define and think about safety can highlight new opportunities for collection and analysis of safety-relevant data. Developing an integrated safety picture to better inform safety-related decision making and policies depends upon identifying, collecting, and interpreting safety-producing behaviors in addition to safety-reducing behaviors. Opportunities and challenges in collecting and analyzing the largely unexploited data on desired, safety-producing operator behaviors are discussed.

Focusing on undesired operator behaviors is pervasive in system design and safety management cultures in aviation. This is evidenced by the extent and range of resources put into eliminating, reducing the likelihood, reducing the consequences, and conducting investigations of adverse states or events. This focus, however, limits the data that are collected, the questions that are asked during data analysis, and therefore our understanding of what operators do in everyday work. Humans play an integral role in aviation safety. Therefore, human performance represents a significant source of aviation safety data. Human performance includes both desired and undesired actions. Most of the time, those actions *promote* safety, but sometimes those actions can *reduce* safety. Commercial aviation hull-loss accidents today are significantly below 1 per million flights, and have been steadily decreasing since the advent of commercial jet operations in 1958. Since as early as 1967, no year has had more than 4 hull losses per million flights (Airbus, 2020). While it is difficult, if not impossible, to tease out the individual contributions of the significant advances that have been made in hardware, software, and human factors to the steady reduction in these accidents over the years, the fact that hull loss accidents were exceedingly rare even in the decades preceding these advances suggests that humans have been making significant contributions to aviation safety throughout its history. When our safety thinking systematically restricts the data we collect and analyze, however, this restricts our opportunities to learn from human performance. Importantly, when this restriction is systematic, it can bias what we learn, which can, in turn, affect our safety policies and decision making.

Most of our learning about human performance in aviation comes from studying relatively rare errors and failures. The magnitude of the discrepancy between human performance that is representative of operations and human performance that is actually analyzed can be difficult to grasp without data to provide some base rate context. For example, it is widely reported that human error has been implicated in 70%-80% of aviation accidents (e.g., Wiegmann & Shappell, 2003). Additionally, analysis of Line Operational Safety Audit (LOSA) data indicates that pilots intervene to manage aircraft malfunctions on 20% of normal flights (PARC/CAST, 2013). If those percentages are examined within the context of 10 years of worldwide jet data (Boeing, 2017), a contingency table can be constructed, depicting Outcome (not accident or accident) by whether human intervention was identified as being associated with that outcome (see Figure 1). When only data from errors and failures are analyzed, the vast majority of human impacts on system performance are largely ignored. Not only does this indicate significant missed opportunities for learning, but it implies that learning is focused on a very small sample of non-representative data, which can have significant impacts on what is learned and how those insights inform safety policies and safety-related decision making.

		Outcome		
		Not Accident	Accident	
Attributed to Human Intervention	No	195,199,690	78	195,199,768
	Yes	48,799,922	310	48,800,232
		243,999,612	388	244,000,000

Figure 1. Human contributions to safety successes (solid oval) far outweigh their contributions to failures (dashed oval), but are relatively unstudied and poorly understood. Note: Accident is defined as hull loss and/or at least one fatality.

While making generalizations from a sample of population data is common practice, responsible researchers work to ensure that their sample is representative of the population of interest. Failure to do so can result in sampling bias, in which what is learned from the sample is erroneously attributed to the larger population. When a sample is systematically non-representative, generalizations from the sample data are suspect. An assertion, for example, that human error contributes to accidents, therefore removing humans will reduce accidents, ignores that humans are also a significant source of successful system performance, and in fact contribute to safety far more than they reduce safety. Indeed, extrapolating from the data in Figure 1 suggests that pilots intervene to keep flights safe over 157,000 times for every time that pilot error contributes to an accident.

### Changing Our Safety Thinking

Changes in how we define and think about safety can highlight new opportunities for collection and analysis of safety-relevant data. Hollnagel (2016) has proposed that we update our definition of safety to include not only minimizing opportunities for undesired states, but also

maximizing opportunities for desired states. This approach is better aligned to understanding human performance, which contributes to both. Furthermore, to maximize opportunities to learn from human performance, data should be collected and analyzed on routine performance, not just exceptional performance. Learning only from rare events means that learning only occurs rarely. While learning from frequent successes has the advantages of increasing sample rate, sensitivity, and timeliness of safety learning, it raises important issues about determining exactly what data to capture, how to analyze and manage this potentially massive expansion of safety data, and translating learned insights into policy and design decisions.

### **Collecting and Analyzing Data on Safety-Producing Behaviors**

Most aviation organizations already collect data on operator performance from various sources, including operator-, observer-, and system-generated data. As we expand our understanding of what constitutes a safety-relevant occurrence to include both desired as well as undesired behaviors, this raises questions about the collection of data on operators' safety-producing behaviors. Can we leverage data that are already being collected, and are there new opportunities for data collection based on our expanded safety thinking? Much of the data we collect are only analyzed for "safety exceedances". While we are actually collecting significant data on "what goes right," even in our safety reporting systems that focus on "what goes wrong," our analysis processes do not often consider these data. Similarly, flight data recorders capture "what happens," including both desired and undesired actions, but our analyses typically focus on the undesired. Thus, there are potential opportunities for data collection (i.e., systematically collecting data on what goes well), as well as data analysis (i.e., analyzing data we have already collected, but may not have previously considered relevant).

#### **Operator-Generated Data**

Operator-generated data include interviews, questionnaires, and event self-reports about an operator's own lived experience. These data represent perhaps the best source of insights into what the operators may have been thinking about, including their motivations, intentions, goals, and pressures, and how they believe those considerations may have influenced their decisions and actions. Although some modes of thought are not open to introspection (see Kahneman, 2011), insights can still be gained from how operators think about their thinking. Similarly, operator-generated data affords an opportunity to learn from how operators talk about their own safety-producing performance. We have well-established shared terminology for describing risks, hazards, and errors (e.g., Wiegmann & Shappell, 2003), but do not yet have such a vocabulary for safety-producing behaviors. Can we identify the parlance already in use for describing safety-producing behaviors, and can we use that to bootstrap development of a new safety language that can address both desired and undesired behaviors? This question is explored elsewhere in these Proceedings by Feldman et al. (2021), using an existing event report collection: NASA's Aviation Safety Reporting System (ASRS). Although ASRS reports are collected for the purpose of describing something that went wrong, they may represent a valuable source of data on things going well, particularly related to noticing, tracking, responding to, and learning from the described problem event (e.g., Holbrook et al., 2020).

#### **Observer-Generated Data**

Observer-generated data include data from observations of line operations as well as training and simulated events. These data represent an excellent source of insight into overt

behaviors, particularly those that may not be salient to operators. This is particularly relevant for safety-producing behaviors that operators may see as “routine” or “just part of the job” and therefore may be less likely to self-report those behaviors. In rich and complex environments like aircraft flight decks, not every behavior is realistically observable – there is simply too much going on, and not every action is meaningful to capture. Knowledge frameworks are often used to train and prepare observers for what to focus on. These knowledge frameworks can be thought of as one way of embodying safety thinking. There are some well-established knowledge frameworks, such as Threat and Error Management (TEM), which is the basis for LOSA (Klinect, Wilhelm, & Helmreich, 1999). While TEM uses undesired behaviors and states as the primary triggers for data collection, this framework still affords opportunities to collect data on how pilots safely managed threats and errors. American Airlines has developed a “Learning and Improvement Team” framework for flight line observations that is explicitly designed for the collection of flight crew resilient performance (American Airlines, 2020). Exploration of how the knowledge frameworks of observers affect the insights they derive from an observation is discussed elsewhere in these Proceedings (Mumaw, Billman, & Holbrook, 2021).

### **System-Generated Data**

System-generated data include flight data records as well as documentation of flight regulations and procedures. Automation enables collecting a large volume of system data on what is happening, via flight data recordings, with less overhead at the time of collection than operator- and observer-generated data. Here, the focus on risks and hazards manifests in terms of which data we choose to analyze – that is, a failure state or adverse event triggers analysis that leaves the vast majority of collected data unconsidered. Indeed, commercial airlines with Flight Operations Quality Assurance (FOQA) programs use data from flight data recorders to monitor daily operations, but often only look at the data from flights with known adverse events (i.e., flights that violate some pre-determined “safety exceedance” criterion). “Non-event” flight data may be analyzed to establish a baseline for comparison, but not as a valuable source of learning, themselves. These “non-event” flights, however, can afford opportunities for insights into safety-producing behaviors, such as actions taken by flight crews to mitigate or prevent adverse events from manifesting (e.g., Holbrook et al., 2019). That is, the occurrence of the adverse event does not have to be a pre-requisite for learning. The amount of flight data collected opens up application of “big data” approaches to analysis, and flight data represent an excellent source of data on quantitative performance parameters, such as timing or frequency. But while flight data can provide many quantitative details about operator and vehicle performance, they cannot provide information about the knowledge state, motivation, or broader context for the event. This information could be obtained through observer- or operator-generated data to supplement system data and provide a more complete understanding of the event.

### **Human-in-the-Loop (HITL) Flight Simulation**

HITL flight simulations represent an additional approach to collecting data on human performance. While data from real-world operations offer the most veridical glimpse into everyday work, HITLs provide opportunities to collect multiple sources of data from the same event, as well as the capability to more efficiently test new approaches to data collection. One of the challenges in designing HITL simulations from a safety mindset focused primarily on errors and failures is that, to have something to measure, scenarios must be designed to induce those errors and failures, which can be difficult when studying high-performing workers. Ironically, in

this situation, the resilient, safety-producing performance of the test participants is seen as an impediment to data collection, rather than important data to be collected – an obstacle that can sometimes require experimenters to create scenarios that are far-removed from representative flight operations in order to induce the errors and failures “required” for performance measurement. Scenarios can be designed, however, to include events and perturbations that might not otherwise involve enough risk or hazard to trigger data collection or analysis in real-world operations, and thus afford opportunities for observation and measurement of more “routine” performance. While it is always a concern that participants in any simulation will not perform in the same manner that they do in actual operations, perhaps this effect may be somewhat mitigated by designing simulation scenarios that are more representative of real-world operations. Exploration of HITL simulation as an approach to learning about pilots’ safety-producing behaviors is discussed elsewhere in these Proceedings (Stephens et al., 2021).

### **Implications for system design**

Changing the way we think about safety is not just relevant to system operators but also to system designers. Our safety thinking affects our design assumptions, which are influenced by our understanding of human performance. We are just beginning, however, to build an understanding of safety-producing behavior. A focus on failure alone can lead to design assumptions about improving safety by minimizing human roles and the need to protect the system from error-prone humans. While we certainly should acknowledge human limitations and the consequences of human error, system designs should also leverage the capabilities of humans to create and sustain safe operations. If these capabilities are poorly understood, what assumptions are going into the design of the increasingly autonomous machine systems to which these functions may be relegated? Challenges and opportunities for system design that leverage human capabilities and new ways of thinking about safety are expanded upon elsewhere in these Proceedings (Lachter, Hobbs, & Holbrook, 2021; Nemeth & Holbrook, 2021).

### **Conclusions**

While we should continue to learn from what goes wrong, we should also try to learn from what goes right. Learning from what goes right can enable us to make data-informed adjustments to operations and policies without having to wait for something to go wrong. Changing how we define safety – expanding our understanding of what constitutes a safety-relevant issue – can inform this learning, and is relevant to both operations and system design. This expansion in thinking brings with it a need to expand methods of data collection and analysis, representing an important opportunity for human factors and human performance communities.

### **Acknowledgments**

This work was funded by NASA’s System-Wide Safety Project, part of the Aeronautics Research Mission Directorate’s Aviation Operations and Safety Program.

### **References**

Airbus. (2020). A statistical analysis of commercial aviation accidents 1958-2019. Reference X00D17008863 Issue 4. France: Art & Caractere.

- American Airlines' Department of Flight Safety (2020). Trailblazers into Safety-II: American Airlines' Learning and Improvement Team, A White Paper Outlining AA's Beginnings of a Safety-II Journey.
- Boeing. (2017). Statistical Summary of Commercial Jet Airplane Accidents: Worldwide Operations 1959 – 2016. Seattle, WA: Aviation Safety, Boeing Commercial Airplanes.
- Feldman, J., Barshi, I., Smith, B., & Matthews, B. (2021). Reports of resilient performance: Investigating operators' descriptions of safety-producing behaviors in the ASRS. In Proceedings of the 2021 International Symposium on Aviation Psychology.
- Holbrook, J., Prinzel, L. J., Stewart, M. J., & Kiggins, D. (2020). How do pilots and controllers manage routine contingencies during RNAV arrivals? In Proceedings of the 11<sup>th</sup> International Conference on Applied Human Factors & Ergonomics.
- Holbrook J. B., Stewart, M. J., Smith, B. E., Prinzel, L. J., Matthews, B. L., Avrekh, I., Cardoza, C. T., Ammann, O. C., Adduru, V., & Null, C. H. (2019). Human performance contributions to safety in commercial aviation. NASA/TM-2019-220417.
- Hollnagel, E. (2016). Resilience Engineering. [erikhollnagel.com/ideas/resilienceengineering.html](http://erikhollnagel.com/ideas/resilienceengineering.html). Accessed March 1, 2021.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus, & Giroux.
- Klinec, J.R., Wilhelm, J.A., Helmreich, R. L. (1999). Threat and Error Management: Data from Line Operations Safety Audits. In Proceedings of the Tenth International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University.
- Lachter, J., Hobbs, A., & Holbrook, J. (2021). Thinking outside the box: The human role in increasingly automated aviation systems. In Proceedings of the 2021 International Symposium on Aviation Psychology.
- Mumaw, R. J., Billman, D., & Holbrook, J. (2021). How do different knowledge frameworks help us learn from aviation line operations? In Proceedings of the 2021 International Symposium on Aviation Psychology.
- Nemeth, C., & Holbrook, J. (2021). Resilience engineering's potential for Advanced Air Mobility. In Proceedings of the 2021 International Symposium on Aviation Psychology.
- PARC/CAST Flight Deck Automation Working Group. (2013). Operational use of flight path management systems. Final Report of the Performance-based operations Aviation Rulemaking Committee/Commercial Aviation Safety Team Flight Deck Automation Working Group. Washington, DC: Federal Aviation Administration.
- Stephens, C., Prinzel, L. J., Kiggins, D., Ballard, K., & Holbrook, J. (2021). Evaluating the use of high-fidelity simulator research methods to study airline flight crew resilience. In Proceedings of the 2021 International Symposium on Aviation Psychology.
- Wiegmann, D. A., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis: The Human Factors Analysis and Classification System*. London, UK: Ashgate.